

日本国特許庁
JAPAN PATENT OFFICE

別紙添付の書類に記載されている事項は下記の出願書類に記載されて
いる事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed
with this Office

出願年月日

Date of Application:

2002年11月11日

出願番号

Application Number:

特願2002-326257

[ST.10/C]:

[JP2002-326257]

出願人

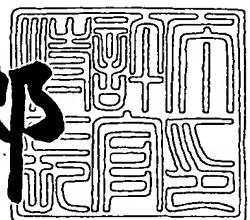
Applicant(s):

株式会社日立製作所

2003年 2月28日

特許庁長官
Commissioner,
Japan Patent Office

太田信一郎



出証番号 出証特2003-3011620

【書類名】 特許願

【整理番号】 K02009451A

【あて先】 特許庁長官殿

【国際特許分類】 G06F 3/06

【発明者】

【住所又は居所】 神奈川県小田原市中里322番地2号 株式会社日立製作所 R A I D システム事業部内

【氏名】 井戸 健嗣

【発明者】

【住所又は居所】 神奈川県足柄上郡中井町境781番地 日立コンピュータ機器株式会社内

【氏名】 柏木 満

【特許出願人】

【識別番号】 000005108

【氏名又は名称】 株式会社日立製作所

【代理人】

【識別番号】 100075096

【弁理士】

【氏名又は名称】 作田 康夫

【手数料の表示】

【予納台帳番号】 013088

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

【物件名】 要約書 1

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 記憶システム

【特許請求の範囲】

【請求項1】 送られた情報を格納する第1の記憶領域と、
送られた情報を格納する第2の記憶領域と、
前記第1の記憶領域を制御する第1の記憶制御装置と、
前記第2の記憶領域を制御する第2の記憶制御装置と、
前記第1の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求に従って、前記第1の記憶領域に対してデータの書き込みを行う一方、前記第2の記憶制御装置に対して前記第1の書き込み要求を転送するものであり、
前記第2の記憶制御装置は、前記第1の記憶制御装置から受けた前記第1の書き込み要求に従って前記第2の記憶領域に対してデータを書き込む動作を行いながら、前記第2の記憶領域に対して送られた第2の書き込み要求を、前記第1の記憶制御装置に対して転送するものであることを特徴とする記憶システム。

【請求項2】 請求項1記載の記憶システムにおいて、

前記第1の記憶制御装置は、前記第2の記憶制御装置から送られた第2の書き込み要求に従って、前記第1の記憶領域に対してデータを書き込む一方、前記第2の記憶制御装置に対して前記第2の書き込み要求を転送するものであることを特徴とする記憶システム。

【請求項3】 請求項2記載の記憶システムにおいて、

前記第2の記憶制御装置は、前記第1の記憶制御装置から送られた第2の書き込み要求に従って、前記第2の記憶領域に対してデータを書き込むものであることを特徴とする記憶システム。

【請求項4】 送られた情報を格納する第1の記憶領域と、

前記第1の記憶領域を制御する第1の記憶制御装置とを有する第1の記憶システム及び、

送られた情報を格納する第2の記憶領域と、

前記第2の記憶領域を制御する第2の記憶制御装置とを有する第2の記憶システムとを有する記憶システム群に利用される第1の記憶システムであって、

前記第1の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求に従って、前記第1の記憶領域に対してデータの書き込みを行う一方、前記第1の記憶制御装置は、前記第2の記憶領域に書き込むための前記第1の書き込み要求を、前記第2の記憶制御装置へ転送する一方、

前記第1の記憶制御装置は、前記第2の記憶領域に対して送られた第2の書き込み要求を受け取ることを特徴とする第1の記憶システム。

【請求項5】 請求項1記載の記憶システムにおいて、

前記第1の記憶制御装置は、前記第2の記憶制御装置から送られた第2の書き込み要求に従って、前記第1の記憶領域に対してデータを書き込む一方、前記第2の記憶制御装置に対して前記第2の書き込み要求を転送するものであることを特徴とする記憶システム。

【請求項6】 請求項2記載の記憶システムにおいて、

前記第2の記憶制御装置は、前記第1の記憶制御装置から送られた第2の書き込み要求に従って、前記第2の記憶領域に対してデータを書き込むものであることを特徴とする記憶システム。

【請求項7】 送られた情報を格納する第1の記憶領域と、

前記第1の記憶領域を制御する第1の記憶制御装置とを有する第1の記憶システム及び、

送られた情報を格納する第2の記憶領域と、

前記第2の記憶領域を制御する第2の記憶制御装置とを有する第2の記憶システムとを有する記憶システム群に利用される第1の記憶システムであって、

前記第2の記憶制御装置は、前記第1のホストコンピュータから送られた第1の書き込み要求を前記第2の記憶領域に書き込む一方、

前記第2の記憶制御装置は、前記第2のホストコンピュータから前記第2の記憶領域に対して送られた第2の書き込み要求を、前記第1の記憶制御装置に対して転送するものであることを特徴とする第2の記憶システム。

【請求項8】 請求項1記載の記憶システムにおいて、

前記第1の記憶制御装置は、前記第2の記憶制御装置から送られた第2の書き込み要求に従って、前記第1の記憶領域に対してデータを書き込む一方、前記第2

の記憶制御装置に対して前記第2の書き込み要求を転送するものであることを特徴とする記憶システム。

【請求項9】 請求項2記載の記憶システムにおいて、

前記第2の記憶制御装置は、前記第1の記憶制御装置から送られた第2の書き込み要求に従って、前記第2の記憶領域に対してデータを書き込むものであることを特徴とする記憶システム。

【請求項10】 送られた情報を格納する第1の記憶領域と、

前記第1の記憶領域を制御する第1の記憶制御装置とを有する第1の記憶システム及び、

送られた情報を格納する第2の記憶領域と、

前記第2の記憶領域を制御する第2の記憶制御装置とを有する第2の記憶システムとを有する記憶システム群に利用されるデータ転送方法であって、

前記第1の記憶制御装置は、

前記第1の記憶領域に対して送られた第1書き込み要求を受け取り、

前記第1の書き込み要求に従って、前記第1の記憶領域に要求された情報を書き込む一方、前記第2の記憶制御装置に対して前記第1の書き込み要求を転送し、

前記第2の記憶制御装置は、

前記第1の記憶装置から送られた前記第1の書き込み要求を受け取り、

前記第1の書き込み要求に従って、前記第2の記憶領域に要求された情報を書き込みつつ、前記第2の記憶領域に対して送られた第2の書き込み要求を受け取り

前記第1の記憶制御装置へ前記第2の書き込み要求を転送することを特徴とするデータ転送方法。

【請求項11】 請求項10記載のデータ転送方法において、

前記第1の記憶制御装置は、前記第2の記憶制御装置から送られた第2の書き込み要求に従って、前記第1の記憶領域に対してデータを書き込む一方、前記第2の記憶制御装置に対して前記第2の書き込み要求を転送することを特徴とするデータ転送方法。

【請求項12】 請求項10記載のデータ転送方法において、

前記第2の記憶制御装置は、前記第1の記憶制御装置から送られた第2の書き込み要求に従って、前記第2の記憶領域に対してデータを書き込むことを特徴とするデータ転送方法。

【請求項13】 送られた情報を格納する第1の記憶領域と、

前記第1の記憶領域を制御する第1の記憶制御装置とを有する第1の記憶システム及び、

送られた情報を格納する第2の記憶領域と、

前記第2の記憶領域を制御する第2の記憶制御装置とを有する第2の記憶システムとを有する記憶システム群に利用されるデータ転送方法であって、

前記第1の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求を受け取り、

前記第1の書き込み要求に従って、前記第1の記憶領域に要求された情報を書き込む一方、転送するための前記第1の書き込み要求を転送つつ、前記第2の記憶領域に対して送られた第2の書き込み要求を前記第1の記憶制御装置へ転送することを特徴とするデータ転送方法。

【請求項14】 請求項13記載のデータ転送方法において、

前記第1の記憶制御装置は、前記第2の記憶制御装置から送られた第2の書き込み要求に従って、前記第1の記憶領域に対してデータを書き込む一方、前記第2の記憶制御装置に対して前記第2の書き込み要求を転送することを特徴とするデータ転送方法。

【請求項15】 請求項14記載のデータ転送方法において、

前記第2の記憶制御装置は、前記第1の記憶制御装置から送られた第2の書き込み要求に従って、前記第2の記憶領域に対してデータを書き込むことを特徴とするデータ転送方法。

【請求項16】 送られた情報を格納する第1の記憶領域と、

前記第1の記憶領域を制御する第1の記憶制御装置とを有する第1の記憶システム及び、

送られた情報を格納する第2の記憶領域と、

前記第2の記憶領域を制御する第2の記憶制御装置とを有する第2の記憶システム

ムとを有する記憶システム群に利用されるデータ転送方法であって、
前記第2の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求を受け取り、

前記第2の記憶領域に書き込む一方、前記第2の記憶領域に対して送られた第2の書き込み要求を受け取り、

前記第1の記憶制御装置に対して転送することを特徴とするデータ転送方法。

【請求項17】 請求項16記載のデータ転送方法において、

前記第1の記憶制御装置は、前記第2の記憶制御装置から送られた第2の書き込み要求に従って、前記第1の記憶領域に対してデータを書き込む一方、前記第2の記憶制御装置に対して前記第2の書き込み要求を転送することを特徴とするデータ転送方法。

【請求項18】 請求項16記載のデータ転送方法において、

前記第2の記憶制御装置は、前記第1の記憶制御装置から送られた第2の書き込み要求に従って、前記第2の記憶領域に対してデータを書き込むことを特徴とするデータ転送方法。

【請求項19】 送られた情報を格納する第1の記憶領域と、

前記第1の記憶領域を制御する第1の記憶制御装置と、

前記第1の記憶制御装置内の第1の接続手段と、

送られた情報を格納する第2の記憶領域と、

前記第2の記憶領域を制御する第2の記憶制御装置と、

前記第2の記憶制御装置内の第2の接続手段と、

前記第1の接続手段と前記第1の記憶領域との間でデータを転送する第1の経路及び、前記第1の記憶領域と前記第2の記憶制御装置との間でデータを転送する第2の経路とを有する前記第1の記憶制御装置であり、

前記第1の記憶制御装置と前記第2の記憶領域との間でデータを転送する第3の経路と、

前記第2の接続手段と前記第1の記憶制御装置との間でデータを転送する第4の経路とを有する前記第2の記憶制御装置であることを特徴とする記憶システム。

【請求項20】 請求項19記載の記憶システムにおいて、

前記第1の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求を前記第1の経路を介して前記第1の記憶領域に書き込む一方、前記第1の書き込み要求を前記第2の経路を介して前記第2の記憶制御装置へ転送するものであることを特徴とする記憶システム。

【請求項21】 請求項19記載の記憶システムにおいて、

前記第2の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求を前記第3の経路を介して前記第2の記憶領域に書き込む一方、前記第2の記憶領域に対して送られた第2の書き込み要求を、前記第4の経路を介して前記第1の記憶制御装置に前記第2の書き込み要求を転送するものであることを特徴とする記憶システム。

【請求項22】 送られた情報を格納する第1の記憶領域と、

前記第1の記憶領域を制御し、第2の記憶制御装置と接続された第1の記憶制御装置と、

前記第1の記憶制御装置内の第1の接続手段と、

前記第1の接続手段と前記第1の記憶領域との間で、データを転送される第1の転送経路と、

前記第1の記憶領域と前記第2の記憶制御装置との間で、データを転送される第2の転送経路とを有することであることを特徴とする第1の記憶システム。

【請求項23】 請求項22記載の記憶システムにおいて、

前記第1の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求を前記第1の経路を介して前記第1の記憶領域に書き込む一方、前記第1の書き込み要求を前記第2の経路を介して前記第2の記憶制御装置へ転送するものであることを特徴とする記憶システム。

【請求項24】 請求項22記載の記憶システムにおいて、

前記第2の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求を前記第2の経路を介して受け取り、前記第2の記憶領域に書き込むものであることを特徴とする記憶システム。

【請求項25】 送られた情報を格納する第1の記憶領域と、

前記第1の記憶領域を制御する第1の記憶制御装置と、

前記第1の記憶制御装置内の第1の接続手段と、
送られた情報を格納する第2の記憶領域と、
前記第2の記憶領域を制御する第2の記憶制御装置と、
前記第2の記憶制御装置内の第2の接続手段とを有する記憶システムであって、
第2の接続手段と前記第1の記憶制御装置との間でデータを転送される第1の転送経路と、
前記第1の記憶制御装置と前記第2の記憶領域との間で、データを転送される第2の転送経路とを有するものであることを特徴とする第2の記憶システム。

【請求項26】 請求項25記載の記憶システムにおいて、
前記第2の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求を前記第2の経路を介して前記第2の記憶領域に書き込む一方、第2の記憶領域に対して送られた第2の書き込み要求を前記第1の経路を介して前記第1の記憶制御装置へ転送するものであることを特徴とする記憶システム。

【請求項27】 請求項25記載の記憶システムにおいて、
前記第2の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求を前記第2の経路を介して受け取り、前記第2の記憶領域に書き込むものであることを特徴とする記憶システム。

【請求項28】 送られた情報を格納する第1の記憶領域と、
前記第1の記憶領域を制御する第1の記憶制御装置と、
前記第1の記憶制御装置内の第1の接続手段と、
送られた情報を格納する第2の記憶領域と、
前記第2の記憶領域を制御する第2の記憶制御装置と、
前記第2の記憶制御装置内の第2の接続手段とを有する記憶システムに利用される第1の記憶制御装置であって、
前記第1の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込み要求を第1の記憶領域に書き込む一方、前記2の記憶領域に対して送られた第2の書き込み要求を受け取り前記第1の記憶領域に書き込むものであることを特徴とする第1の記憶制御装置。

【請求項29】 送られた情報を格納する第1の記憶領域と、

前記第1の記憶領域を制御する第1の記憶制御装置と、
送られた情報を格納する第2の記憶領域と、
前記第2の記憶領域を制御する第2の記憶制御装置とを有する記憶システムに利
用される第2の記憶制御装置であって、
前記第2の記憶制御装置は、前記第1の記憶領域に対して送られた第1の書き込
み要求を受け取り、前記第2の記憶領域に書き込む一方、前記第2の記憶領域に
に対して送られた第2の書き込み要求を受け取り、前記第2の書き込み要求を前記
第1の記憶制御装置に転送するものであることを特徴とする第2の記憶制御装置

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、複数の記憶システムのそれぞれに設けられた記憶ボリューム間でペ
アを形成しデータを多重化する記憶システムに関して、特に、データの整合性を
確保しつつ、正ホストコンピュータ及び副ホストコンピュータからのアクセス要
求を受ける技術に関する。

【0002】

【従来の技術】

近年の記憶システムでは、記憶制御装置及び記憶ボリュームに対する障害や災
害時に、実行していたオンライン業務等の処理を速やかに回復できるように、記
憶制御装置や記憶制御装置に接続される記憶ボリュームを2重化することがしば
しばある。また、記憶システムの一部または全体を2重化構成にしておくことは
、システムの保守作業を行う場合にも有効である。

【0003】

記憶ボリュームを2重化するための構成の一つとして、2台のホストコンピュ
ータのそれぞれに接続された記憶制御装置の各々は、正ボリューム又は、副ボリ
ュームに接続される。正ホストコンピュータが正ボリュームに対して書き込み処
理を行うと、記憶制御装置は正ボリュームから副ボリュームへデータをコピーす
る。このコピー処理による正ボリュームと副ボリュームとの状態をペア状態と呼

ぶ。ペア状態においては、副ホストコンピュータは、副ボリュームにアクセスできない。

【0004】

【特許文献1】

特開2001-273242号公報

【0005】

【発明が解決しようとする課題】

従来、ペア状態を解除しなければ、副ホストコンピュータが目的の記憶ボリュームのデータにアクセスすることができなかった。つまり、ペア状態を保持したまま、副ホストコンピュータが目的の記憶ボリュームのデータにアクセスすることができなかった。

【0006】

本発明の目的は、2重化された記憶システムにおいて、正ボリュームと副ボリュームに格納されているデータの整合性を保証しながら、正ホストコンピュータと副ホストコンピュータが、あたかも同一の記憶ボリュームを共有しているかの如く、I/O処理を行うことである。

【0007】

さらに、正側と副側の区別を持たずに両側のホストコンピュータがあたかも同一の記憶ボリュームを共有するかの如く、I/O処理を行うことを可能にすることである。

【0008】

【課題を解決するための手段】

上記目的を達成するために、本発明は、副ボリュームに対して通信用ボリュームを設ける。副記憶制御装置は、副ホストコンピュータが発行した書き込み要求または、読み込み要求であるI/O処理を、通信用ボリュームから正側の記憶ボリュームである正ボリュームへのI/O処理として正側の記憶制御装置へ転送する。正側の記憶制御装置は、転送されたI/O処理を、正側のホストコンピュータから発行されたI/O処理と同様に、副ボリュームへデータを転送する。

【0009】

【発明の実施の形態】

図18は、本発明の原理図である。図18の記憶システムにおいて、正側のホストコンピュータ101が正ボリューム601に対してI/O処理を行うと、正側の記憶制御装置301は、ホストコンピュータ101からI/O処理を受け取り、から正ボリューム601に対してI/O処理を行う(a)。

【0010】

また、正側の記憶制御装置は、正側のホストコンピュータ101からのI/O処理を副側の記憶制御装置302へ転送する。正側の記憶制御装置301からI/O処理を受領した副側の記憶制御装置302は、副ボリューム602にI/O処理を行う(b)。

【0011】

一方、副側のホストコンピュータ102から副ボリューム602にI/O処理を行うと、副側の記憶制御装置302は、ホストコンピュータ102からI/O処理を受け取り、I/O処理を正側の記憶制御装置301へ転送する。副側の記憶制御装置302からI/O処理を受け取った正側の記憶制御装置301は、正ボリューム601にI/O処理を行う(c)。

【0012】

さらに、正側の記憶制御装置301は、受け取ったI/O処理を副側の記憶制御装置302に送り戻す。送り戻されたI/O処理を受け取った副側の記憶制御装置302は、副ボリューム602に対して送り戻されたI/O処理を行う(d)。

【0013】

尚、本発明の正ボリューム601、副ボリューム602及び通信用ボリューム603の各々のボリュームは、実際には、一つ以上の記憶媒体で構成されるものである。記憶媒体の例として、磁気ディスクや光ディスク等がある。特に、RAID (Redundant Array for Inexpensive Disks) 方式を採用する場合、正/副記憶制御装置は、ホストコンピュータから論理ボリュームに対して送られたI/O処理を、論理ボリュームに対応する複数のディスクに対してI/O処理を実行する。

【0014】

図1は図18の原理図を詳細に記載した本発明における記憶システムの第1の実施の形態を示した図である。図1の記憶システムにおいて、記憶制御装置301は、チャネルパス201を介してホストコンピュータ101に接続される。記憶制御装置301は、ホストコンピュータ101から要求されるI/O処理を実行する。記憶制御装置301は、チャネルポート401～403、I/O処理制御部（プロセッサ）501～503とそれらの各々に備わる個別メモリ504～506、チャネルポートを介してI/O処理を行うI/O処理制御部（I/O処理制御部）501～503からアクセス可能な共有メモリとキャッシュメモリ507、ドライブポート526、正ボリューム601に対して書き込み処理及び読み込み処理を行うI/O処理制御部522及びこれに備わる個別メモリ524を有する。ドライブポート526は、正ボリューム601と接続される。

【0015】

記憶制御装置302は、記憶制御装置301と同様の構成をとっている。記憶制御装置302は、チャネルパス204を介してホストコンピュータ102に接続され、ホストコンピュータ102から要求されるI/O処理を制御する。記憶制御装置302は、記憶制御装置301と同様の構成であるが、異なる点は、通信用ボリューム603を有することである。

【0016】

通信用ボリューム603は、ホストコンピュータ102からのI/O処理を記憶制御装置301に転送する際に使用する。

【0017】

記憶制御装置301及び記憶制御装置302はチャネルポート402とチャネルポート405間をチャネルパス202によって、チャネルポート403とチャネルポート404間をチャネルパス203によってそれぞれ接続され、この接続により、正ボリューム601と副ボリューム602をペアとして形成している。チャネルパス202、203は、一方向通行路である。本実施の形態において、チャネルパス202、203は、一方向通信路を用いたが、パスを一つにして、ペアを形成しているので、ホストコンピュータ101から正ボリューム601

に対してI/O処理が発行されると、I/O処理を受領した記憶制御装置301は、チャネルパス202及び記憶制御装置302を介して、ペアである副ボリューム602へI/O処理を転送する。

【0018】

一方、ホストコンピュータ102から副ボリューム602に対してI/O処理が発行されると、I/O処理を受領した記憶制御装置302のI/O処理制御部523は、副ボリューム602に対してI/O処理を実行せずに、最初に通信用ボリューム603の処理として制御を変更する。

【0019】

その後、記憶制御装置302は通信用ボリューム603から正ボリューム601へのI/O処理として、チャネルパス203を介して記憶制御装置301へI/O処理を転送する。記憶制御装置302からI/O処理を受け取った記憶制御装置301のI/O処理制御部503は、正ボリューム601へI/O処理を行う。その後、今度は副ボリューム602への処理として、チャネルパス202を経由して記憶制御装置302へI/O処理を更に転送する。記憶制御装置302は、副ボリューム602に対して転送されたI/O処理を行う。

【0020】

図2は、記憶制御装置302の共有メモリ/キャッシュメモリ514内の管理情報を示したものである。ホストコンピュータ102が副ボリューム602へI/O処理を発行する際、I/O処理制御部510は、共有メモリ/キャッシュメモリ514内の通信ボリューム用管理エリア515に、副ボリューム番号516、コマンド519とパラメータ情報520を格納する。

【0021】

ホストコンピュータ102からの書き込み要求の際、I/O処理制御部510は、共有メモリ/キャッシュメモリ514内の受領データ格納領域521へホストコンピュータ102からのデータを格納する。その後、I/O処理制御部510は、通信用ボリューム603へ処理を移すために、I/O処理制御部508へ制御を移す。制御を移されたI/O処理制御部508は、副ボリューム番号516から、対応するペア正ボリューム番号517とパス情報518を求め、I/O

処理の転送先を算出する。その後、I/O処理制御部508は、通信用ボリューム603から正ボリューム601へのI/Oとして、チャネルパス203経由で記憶制御装置301のI/O処理制御部503へコマンド519を発行し、パラメータ情報521を転送する。書き込み処理を伴う場合には、I/O処理制御部508は、ホストコンピュータ102から送られたデータを格納した受領データ格納領域521から格納したデータも転送する。

【0022】

一方、I/O処理制御部508、509、510の処理とは別に、I/O処理制御部523は、受領データ格納領域521に書き込みデータがあれば、書き込みのデータを通信ボリューム603に書き込む。

【0023】

図3～図7は、図1及び図2における処理フローである。フローは、対象ボリュームが正ボリューム601であるか副ボリューム602であるか、またはホストコンピュータ101からのI/Oかホストコンピュータ102からのI/Oかによってフローが異なる。I/Oには、書き込みと読み込みがあるが、書き込みのI/Oの場合、記憶制御装置302は、ホストコンピュータ102からのI/Oであれば、記憶制御装置301に転送する。また、記憶制御装置301からのI/Oであれば、記憶制御装置302は、副ボリューム602に書き込む。

【0024】

図3でI/O処理制御部510(JOB1)の処理を説明する。ホストコンピュータ102から記憶制御装置302へI/Oが発行され、I/O処理制御部510は、コマンド解析(ステップ701)を行う。その後、I/O処理制御部510は、データ転送が可能である状態かどうか、つまり、ペア状態が正常であるか異常であるかを判断する(ステップ702)。ペア状態が異常の時は、正/副ボリュームとのデータの整合性が取れないので、I/O処理制御部510は、ホストコンピュータ102へペア状態が異常であることを報告するチェック報告を行う(ステップ703)。ペア状態が正常ならば、I/O処理制御部510は、通信ボリュームのJOB3へI/O処理の要求を行い(ステップ704)、I/O処理制御部510(JOB1)は、待機状態となる(ステップ705)。通信

ボリュームのJOB3の処理が終了した時点で処理が再開される（ステップ706）。I/O処理制御部510は、ホストコンピュータ102へ終了報告をする（ステップ707）。

【0025】

次に図5でI/O処理制御部508(JOB3)の処理を説明する。I/O処理制御部508がI/O処理制御部510(JOB1)からI/O処理の要求を受信すると、I/O処理制御部508(JOB3)は、正ボリューム情報517、バス情報518を設定し（ステップ713）、ステップ713で準備した情報に従って、記憶制御装置301へコマンドを発行する（ステップ714）。これにより、I/O処理制御部503(JOB4)の処理が起動される。I/O処理制御部508は、パラメータ転送に続き、データ転送を開始（ステップ715）する。転送が終了すれば、I/O処理制御部508(JOB3)は、待機状態になる（ステップ716）。

【0026】

データ転送が終了した時点で、記憶制御装置301のI/O処理制御部503は、I/O処理制御部508に対してJOB4の終了通知をする。これによりI/O処理制御部508(JOB3)は、再起動され（ステップ717）、I/O処理制御部510(JOB1)へ終了通知をする（ステップ718）。

【0027】

図6では、I/O処理制御部508(JOB3)からのI/O処理要求によって起動されるI/O処理制御部503(JOB4)の処理を説明する。まず、I/O処理制御部503は、I/O処理要求によって起動されたJOB4で、コマンド解析（ステップ719）を行う。その後、I/O処理制御部503は、個別メモリ506にコマンドとパラメータを退避する（ステップ720）。尚、本実施の形態では、個別メモリにコマンドとパラメータを退避したが、共有メモリ／キャッシュメモリ507に退避することもできる。その後、I/O処理制御部503は、共有メモリ／キャッシュメモリ507にデータ転送を行う（ステップ721）。つまり、書き込み処理の際、転送データを伴うので、転送データは、I/O処理制御部503によって、共有メモリ／キャッシュメモリ507に格納さ

れる。

【0028】

その後、I/O処理制御部503 (JOB4) は、I/O処理制御部502 (JOB5) に対してI/O処理要求をし（ステップ722）、I/O処理制御部503 (JOB4) は一旦待機状態になる（ステップ723）。I/O処理制御部502 (JOB5) からの終了通知でJOB4の処理が再開され、I/O処理制御部503は、I/O処理制御部508 (JOB3) へ終了通知をして処理を終了する。また、共有メモリ/キャッシュメモリ507に格納された転送データは、I/O処理制御部522によって、正ボリューム601に書き込まれる。

【0029】

更に図7では、I/O処理制御部503 (JOB4) からのI/O処理要求によって起動されるI/O処理制御部502 (JOB5) の処理を説明する。まず、I/O処理制御部502は、副ボリューム番号とパス情報を設定し（ステップ726）、副ボリューム602のI/O処理を行うJOB2へコマンドを発行する（ステップ727）。I/O処理制御部502は、続いてパラメータ/データ転送を行う（ステップ728）。転送が終了したら、I/O処理制御部502は、I/O処理制御部503 (JOB4) へ処理の終了通知を行う（ステップ729）。

【0030】

ステップ728において、転送されるデータが、記憶制御装置302の共有メモリ/キャッシュメモリ514に存在する時は、データの転送を省略することができる。これにより、データの転送を省略できた場合、ホストコンピュータへのレスポンスを向上させることができる。

【0031】

最後に図4では、I/O処理制御部502 (JOB5) からのI/O要求によって起動されるI/O処理制御部509 (JOB2) の処理を説明する。JOB2が起動されると、I/O処理制御部509は、コマンド解析（ステップ708）を行う。その後、I/O処理制御部509は、ペア状態をチェックする（ステップ709）。ペア状態が異常の時、正/副ボリュームの整合性が取れないため

、 I/O処理制御部509は、 I/O処理制御部502 (JOB5) ヘチェック報告を行う (ステップ710)。ペア状態が正常ならば、 I/O処理制御部509は、 続いてコマンド、 パラメータ及びデータの転送処理を行う (ステップ711)。終了したら I/O処理制御部509は、 正側の記憶制御装置301の I/O処理制御部502 (JOB5) ヘ終了通知をして (ステップ712) 処理を終了する。

【0032】

尚、 ステップ711において、 転送されるデータが、 記憶制御装置302の共有メモリ/キャッシュメモリ514に存在する時は、 データの転送を省略することができる。

【0033】

図8は図3～図7の制御方式を用いて処理を行う際の、 副ボリューム602、 通信ボリューム603、 正ボリューム601で起動するJOB間のシーケンスの流れを示した図である。図8によれば、 ホストコンピュータ102からI/O要求を受領したI/O処理制御部510 (JOB1) は、 コマンドとパラメータの受領及び退避や、 ホストからのデータ受領及びキャッシュメモリへのデータ退避などの前処理を行う。その後、 I/O処理制御部510、 通信ボリューム(603) のI/O処理制御部508 (JOB3) へ処理の要求を行い、 I/O処理制御部510 (JOB1) は待機状態となる。I/O処理制御部508 (JOB3) が起動されると、 正ボリューム601のI/O処理制御部503 (JOB4) に対してコマンドを発行する。I/O処理制御部503において、 JOB4が起動されると、 コマンド/パラメータの受領/退避や、 副ボリューム602からのデータ受領などの前処理を行う。その後、 今度は逆に副ボリューム602へデータをコピーするためにI/O処理制御部502は、 I/O処理制御部509 (JOB2) に対してコマンドを発行する。I/O処理制御部509 (JOB2) は、 データの転送が完了したら I/O処理制御部503 (JOB4) に終了通知を行う。終了通知を受領したI/O処理制御部503 (JOB4) は、 I/O処理制御部508 (JOB3) へ終了通知を行う。

【0034】

続いてI/O処理制御部508 (JOB3) は、I/O処理制御部510 (JOB1) に終了通知を行い、I/O処理制御部510 (JOB1) はホストコンピュータ102へ終了報告をする。この一連の処理を行うことで正/副ボリュームのデータの整合性を保ちながらI/O処理を完結させることが可能となる。

【0035】

本実施の形態において読み込みのI/O処理についても、同様の処理で行うが、以下の処理を行えば、副ボリューム602から読み込み可能である。ホストコンピュータ101から正ボリューム601に対して書き込み要求があることを記憶制御装置302に通知し、共有メモリ514に正ボリューム601に対する書き込み要求を格納し、その後、正ボリューム601に書き込む。ホストコンピュータ102から副ボリューム602に対して直接読み込むことができる。読み込み要求によって読み込まれる領域と共有メモリ514に格納された書き込み要求のある書き込み領域が重ならない場合、副ボリューム602から読み込む。一方、読み込み領域が書き込み領域と重なる場合、正ボリューム601に対して読み込み要求を転送し読み込む。

【0036】

第1の実施の形態は、通信ボリュームを定義しているので、記憶制御装置302の各I/O処理制御部は、通信ボリュームと副ボリュームとの処理を判別する必要がなく、I/O処理制御部の処理が単純になる。また、ホストコンピュータ102のユーザは、正ボリュームと副ボリュームとのデータの整合性が取れているため、副ボリュームに対して正ボリュームと区別することなくI/O処理を行うことができる。

【0037】

本実施の形態は、データの転送処理を非同期に行うこともできる。ここで、データの転送処理の非同期とは、ホストコンピュータ102からの書き込み要求のI/Oと、副ボリュームへのデータ転送処理を非同期に処理することである。第2の実施の形態は、図1と同様の形態をとるが、一部の処理が異なる。

【0038】

図9は、データの転送処理を非同期に行う各I/O処理制御部のシーケンス図

である。図8との違いは、I/O処理制御部510（JOB1）がI/O処理制御部508（JOB3）へ通信ボリュームへの処理通知を行う前に、ホストホストコンピュータ102に対して終了報告を送信することである。図9の場合は、図8の処理よりも、ホストコンピュータ102へのレスポンスが向上する。

【0039】

図10は、本発明の第2の実施の形態を示した図である。本実施の形態は、図1における通信ボリューム603の機能を副ボリューム602が兼ねている。図1の場合と異なる動作は、副ボリューム602が通信ボリューム603で動作するJOBにI/O処理を転送する代わりに、副ボリューム602内の別JOBにI/O処理要求を行う点である。本実施の形態の特徴は、図1の場合には記憶ボリュームとしての役割を持たない通信ボリュームを定義しなければならない。しかし、図10の場合には、通信ボリューム603の定義が不要なところである。第3の実施の形態は、通信ボリューム603を定義する必要がないため、副ボリューム602の容量は、通信ボリューム603の分だけ増やすことができる。

【0040】

図11～図15は、図10における構成の制御フローを記述したものである。副ボリューム602のI/O処理を、同一副ボリューム602内の別JOBを使用して正ボリューム601へ転送し、転送が終了した時点で副ボリューム602のJOBを経由してホストコンピュータへ終了報告をするための制御フローを記述したものである。

【0041】

処理の流れは、図3～図7と同様に行われるが、相違点は、以下の通りである。第1の相違点は、正ボリュームからのI/Oでなければ、同一副ボリューム602の別JOBへ正ボリュームへのI/O処理転送要求を行う（ステップ804）点である。第2の相違点は、副ボリューム602でのI/O処理の引継ぎと判断した際（ステップ813）、通信ボリュームで起動されたJOBの場合と同様に、情報設定（ステップ814及び815）を行った後、正ボリューム601へコマンド発行（ステップ816）、パラメータ、データ転送する点（ステップ817）である。

【0042】

図16は、図11～図15の制御方式を用いて処理を行う際の、副ボリューム602、正ボリューム601で起動するJOB間のシーケンスの流れを示した図である。図16の流れは、図8の流れとほぼ同一である。違いは、通信ボリューム603で起動していたJOBの代わりに、副ボリューム602内でJOB3を割当てたところである。

【0043】

本実施の形態においても、非同期の転送を行うことができる。図9の通信ボリュームを副ボリュームに置き換えることによって、データを非同期に転送することができる。

【0044】

また、本実施の形態において、副ボリューム602でホストコンピュータ102からのデータを共有メモリ／キャッシュメモリ514上に一時的に格納しておけば、正ボリューム601からデータコピーの指示が来た時にデータ転送を省略することが可能である。

【0045】

また、通信ボリューム603を副ボリューム602でなく共有メモリ／キャッシュメモリ514に定義することも可能である。共有メモリ／キャッシュメモリ514に通信ボリューム603を定義すれば、副ボリューム602へ書き込む処理がなくなるので、結果としてホストコンピュータへのレスポンスの向上をさせることができる。

【0046】

第3の実施の形態は、正ボリュームと副ボリュームの区別をなくした形態である。正ボリュームと副ボリュームの区別をなくし、両方ともに正ボリュームの位置付けとし、ホストコンピュータ101及びホストコンピュータ102、もしくは記憶制御装置301及び記憶制御装置302間で正ボリュームと副ボリュームを他のホストI/Oから排他した方に正ボリュームの動作とする。一方、排他をしていない方を副ボリューム602の動作として処理をさせることでデータの整合性を保ちながら、両ホストコンピュータ101及びホストコンピュータ102

からI/O処理を行うことを可能にする。その時点での正ボリュームの動作をするボリュームと副ボリュームの動作をするボリュームの関係が決定すれば、処理フローは図3~7、または図11~15の動作に従うことができる。

【0047】

図17は、上記正ボリュームと副ボリュームの区別をなくした本実施の形態における処理を行うフロー図である。あらかじめ、対のボリュームに対して排他モードを定義しておく。即ち一方のボリュームを排他メイン、他方を排他サブとし、メインの排他を取得した記憶制御装置が優先となる。

【0048】

まずホストコンピュータからのI/Oに対するJOBが起動され（ステップ856）、コマンド解析を行う（ステップ857）。次に自ボリュームの排他モードをチェックし（ステップ858）、排他サブであれば続いて他ホストがリザーブ中かどうかチェックする（ステップ859）。リザーブ中であればホストヘリザーブコンフリクト報告を行い（ステップ834）、処理を終了する（ステップ835）。リザーブ中でなければ更に他ホストでボリュームが排他中かどうかをチェックする（ステップ829）。排他中であれば他ホストがアクセス中であるため、一旦JOBは待機状態となる（ステップ836）。排他中でなければ、排他サブを取得する（ステップ830）。

【0049】

その後、相手ボリュームへホストからのパラメータまたはデータを転送する（ステップ831）。この時未だ排他メインを取得しているわけではないので、転送結果をチェックする必要がある。転送結果が相手ボリュームでボリューム排他中であれば（ステップ832）、既に他ホストに排他メインを取得されていたことになるため、一旦排他サブを解除し（ステップ851）、JOBは待機状態になる（ステップ833）。転送結果が「相手ボリュームでリザーブ中」であれば（ステップ837）、すれ違いで他ホストにリザーブを取得されていたことになるため、排他サブを解除し（ステップ852）、ホストヘリザーブコンフリクト応答を行い（ステップ838）、処理を終了する（ステップ839）。転送結果が排他中でもリザーブ中でもなければ、正常に相手ボリュームへ転送処理を行え

ため、自ボリュームに対してもホストからのパラメータまたはデータの転送を行い（ステップ853）、排他サブ情報を解除し（ステップ854）ホストへ終了報告を行い（ステップ840）、処理を終了する（ステップ841）。

【0050】

自ボリュームの排他モードがメインの時にも、まず他ホストがリザーブ中かどうかチェックする（ステップ842）。リザーブ中であればホストへリザーブコンフリクト応答を行い（ステップ848）、処理を終了する（ステップ849）。リザーブ中でなければ更に他ホストでボリュームが排他中かどうかをチェックする（ステップ843）。排他中であれば他ホストがアクセス中であるため、一旦JOBは待機状態となる（ステップ850）。排他中でなければ、排他メインを取得する（ステップ844）。その後、自ボリュームと相手ボリュームへホストからのパラメータまたはデータを転送する（ステップ845）。この時は排他メインを取得しているので転送は必ず終了する。換言すると、すれ違いで相手のボリュームに排他がかかっていても、自ボリュームの排他モードがメインなので、時間監視を行い、相手のボリュームの排他が解除されるのを待ち処理を実行する。その後、排他メイン情報を解除し（ステップ855）ホストへ終了報告を行い（ステップ846）、処理を終了する（ステップ847）。

【0051】

本実施の形態によると、排他メインを取得すれば、ペアのボリュームとのデータの整合性が取れている。よって、読み込み処理の場合、常に正側の記憶制御装置は、自ボリュームからデータ転送を行うことができる。自ボリュームから読み出すことができるので、ペアボリュームへのデータ転送処理のオーバヘッドは、なく、ホストコンピュータへのレスポンス向上をさせることができる。また、正／副ボリュームの区別がないため、クラスタリングに適した形態となる。

【0052】

本発明の第1～第3の実施の形態において、図1～8で、正ボリューム601、副ボリューム602及び通信ボリューム603にI/O処理を行う際、I/O処理制御部522及びI/O処理制御部523が処理を行っているが、このI/O処理をI/O処理制御部501～503、508～510に行わせることもで

きる。

【0053】

また、チャネルパス202、203は、本実施の形態においては専用線を用いたが、専用線だけに限らず、公衆回線、LAN、又は、インターネット等の回線を使用することもできる。

【0054】

また、図1において、正ボリューム601とは別に、通信ボリューム603を記憶制御装置301に接続すれば、ホストコンピュータ101とホストコンピュータ102、記憶制御装置301と記憶制御装置302、正ボリューム601と副ボリューム602の機能を入れ替えることができる。同様に、第2の実施の形態においても、機能を入れ替えることができる。

【0055】

【発明の効果】

本発明によれば、ペア状態であっても、副側のホストコンピュータは、正ボリュームとのデータの整合性をとりながら、副ボリュームにI/O処理を行うことが可能である。

【図面の簡単な説明】

【図1】 本発明の第1の実施の形態における記憶システムの構成例を示す図である。

【図2】 I/O処理を転送するために必要な情報を格納している共有メモリ/キャッシュメモリを示す図である。

【図3】 第1の実施の形態におけるI/O処理制御部510のフロー図である

【図4】 第1の実施の形態におけるI/O処理制御部509のフロー図である

【図5】 第1の実施の形態におけるI/O処理制御部508のフロー図である

【図6】 第1の実施の形態におけるI/O処理制御部503のフロー図である

【図7】 第1の実施の形態におけるI/O処理制御部502のフロー図である。

【図8】 第1の実施の形態における各I/O処理制御部のJOBの関係を示したシーケンス図である。

【図9】 第2の実施の形態における各I/O処理制御部のJOBの関係を示したシーケンス図である。

【図10】 本発明の第3の実施の形態における記憶システムの構成例を示す図である。

【図11】 第3の実施の形態におけるI/O処理制御部510のフロー図である。

【図12】 第3の実施の形態におけるI/O処理制御部509のフロー図である。

【図13】 第3の実施の形態におけるI/O処理制御部508のフロー図である。

【図14】 第3の実施の形態におけるI/O処理制御部503のフロー図である。

【図15】 第3の実施の形態におけるI/O処理制御部502のフロー図である。

【図16】 第3の実施の形態における各I/O処理制御部のJOBの関係を示したシーケンス図である。

【図17】 第4の実施の形態におけるフロー図である。

【図18】 本発明の動作原理を示す原理図である。

【符号の説明】

101、102・・・ホストコンピュータ、

201、202、203、204・・・チャネルバス、

301、302・・・記憶制御装置、

501、502、503、508、509、510、522、523・・・I/O処理制御部（プロセッサ）、

507、514・・・I/O処理制御部間共有メモリ／キャッシュメモリ、

601、602、603・・・記憶装置、

515・・・通信ボリューム用管理エリア、

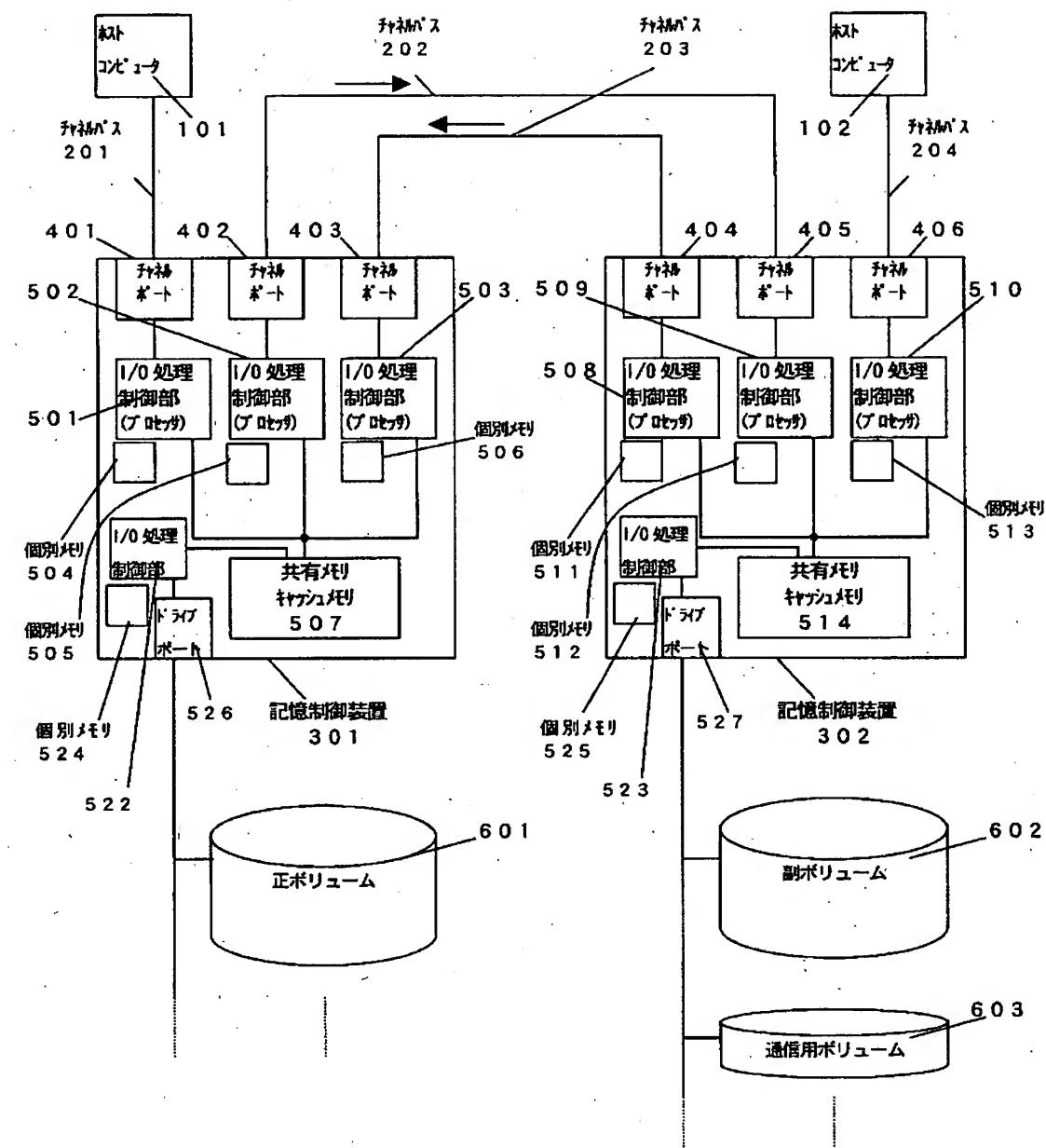
516・・・副ボリューム番号情報、

517・・・ペア正ボリューム番号情報

【書類名】 図面

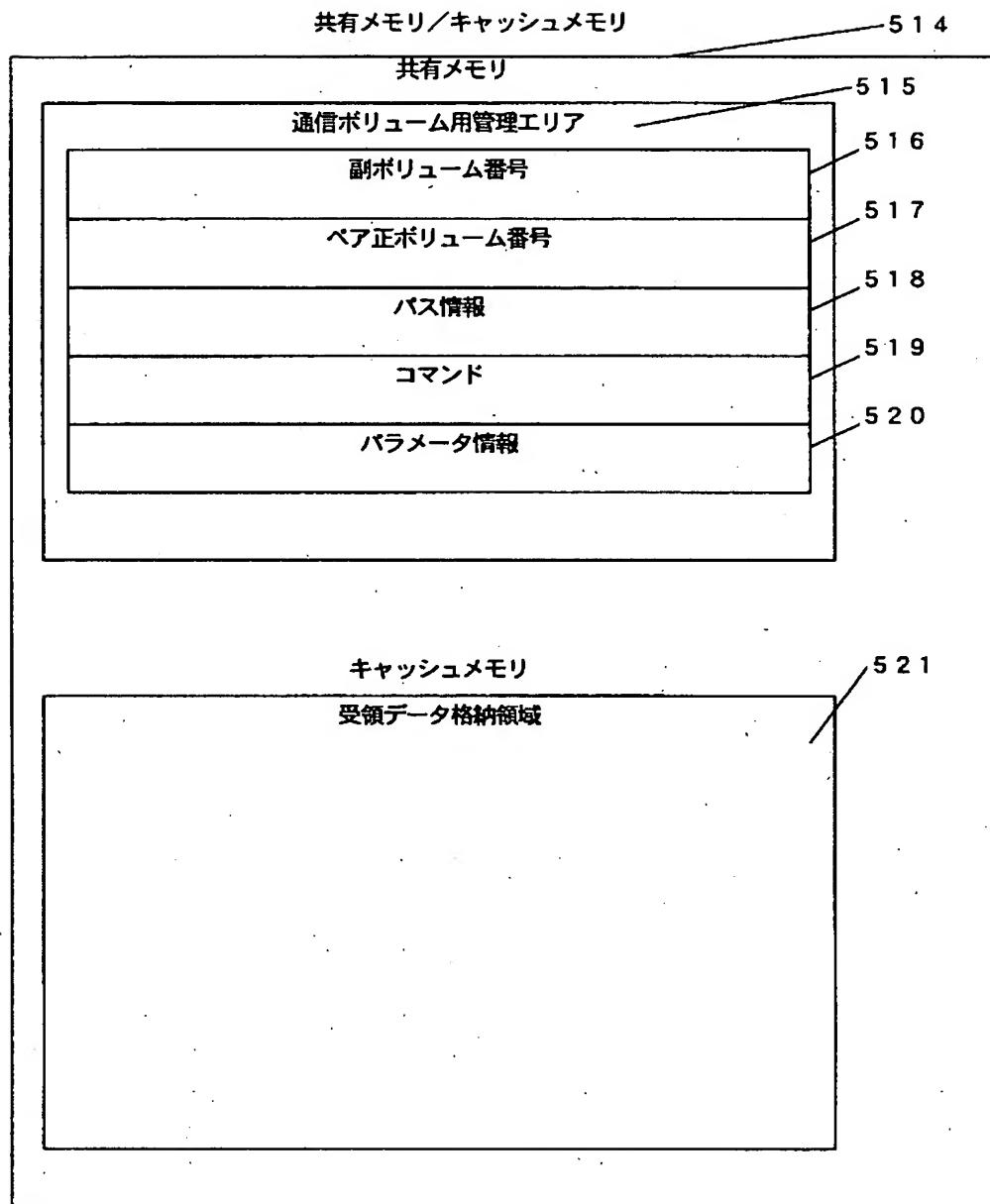
【図1】

図1



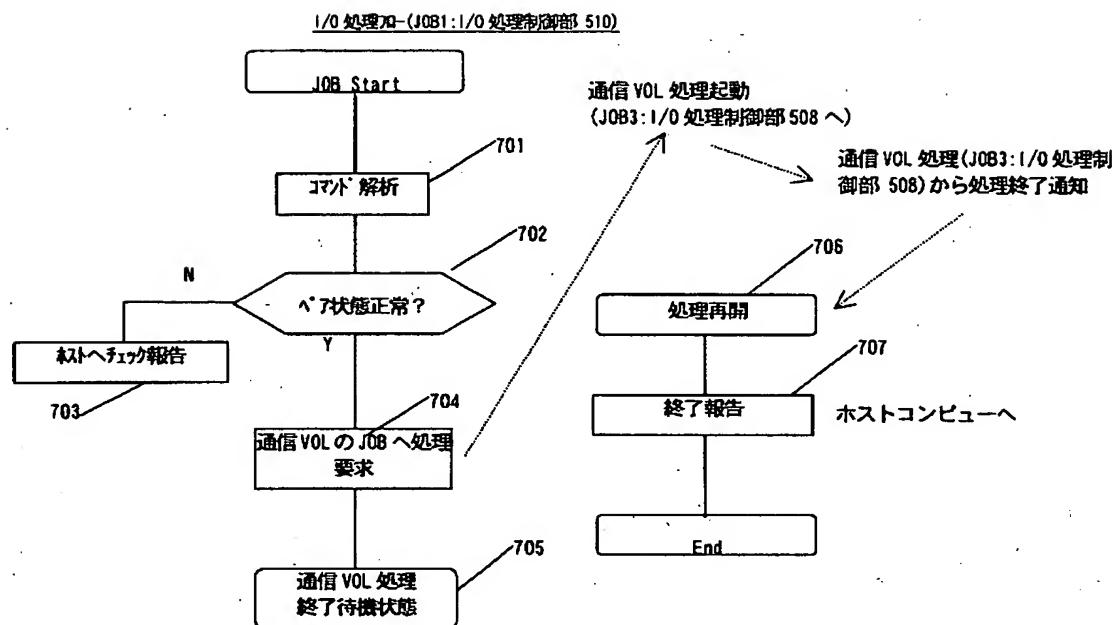
【図2】

図2

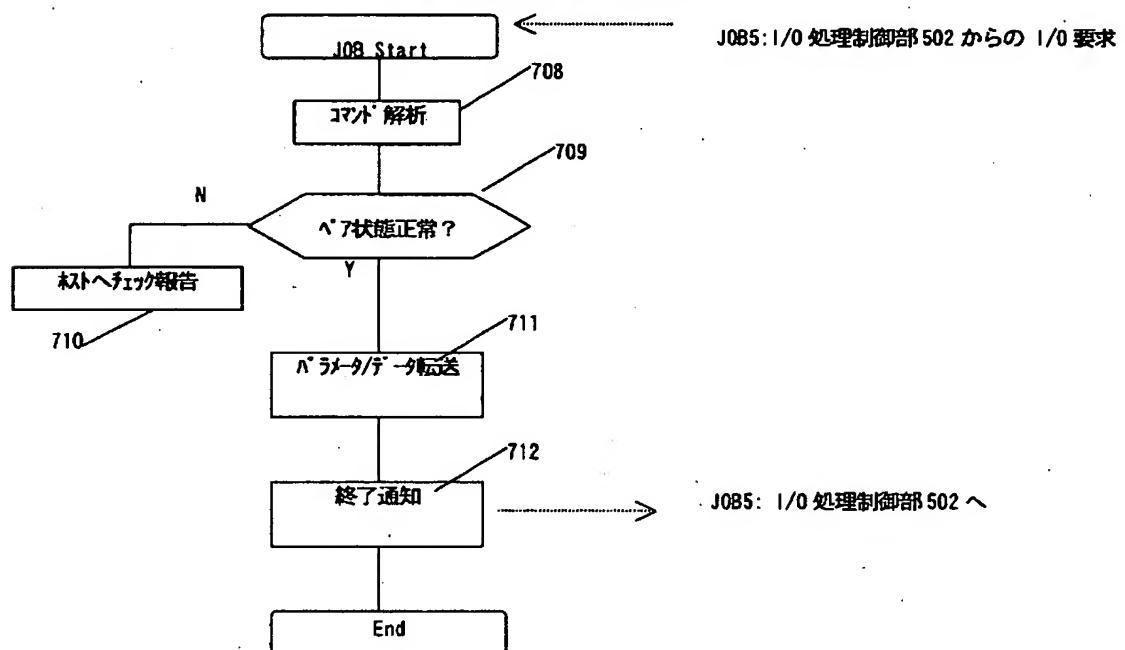


【図 3】

図 3

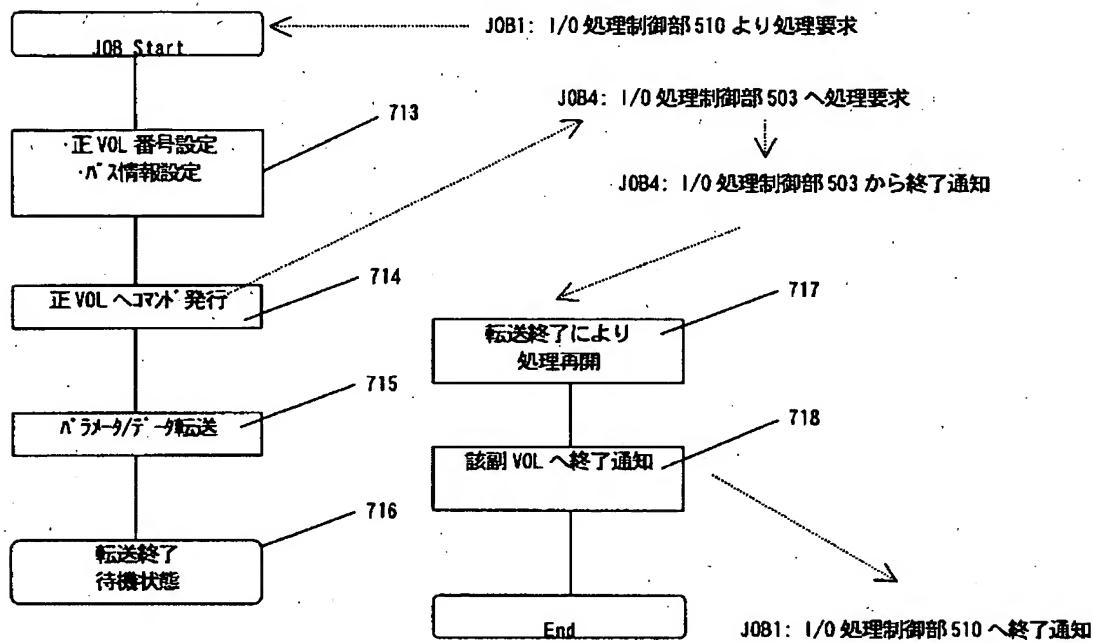


【図 4】

図 4
I/O 处理加-(JOB2:I/O 处理制御部 509)

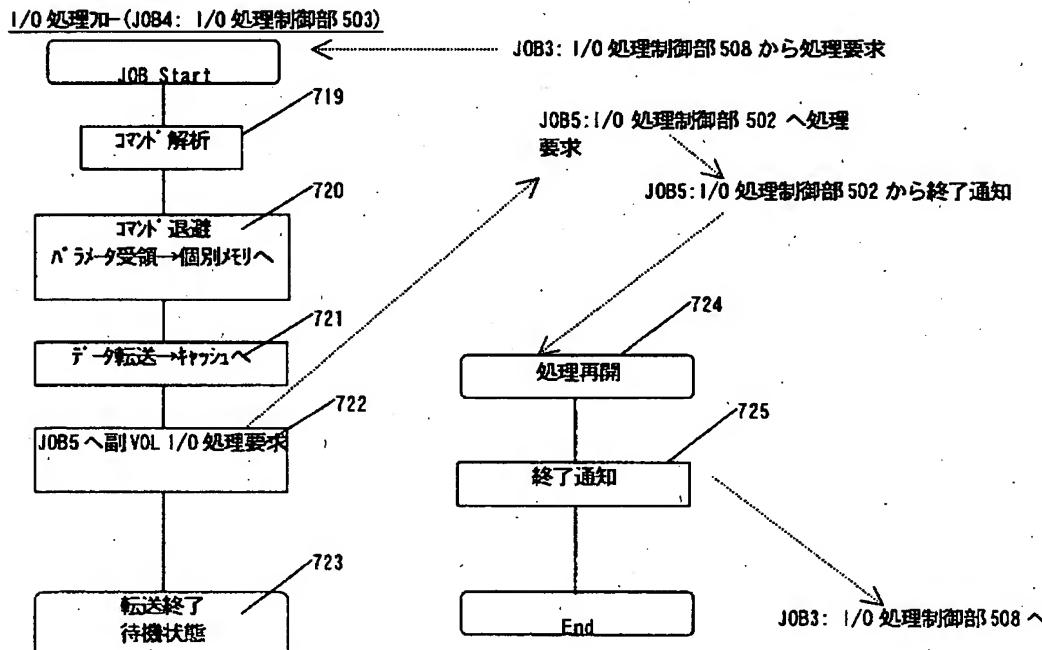
【図5】

図5

通信 VOL I/O 处理加-JOB3:I/O 处理制御部 508 の処理

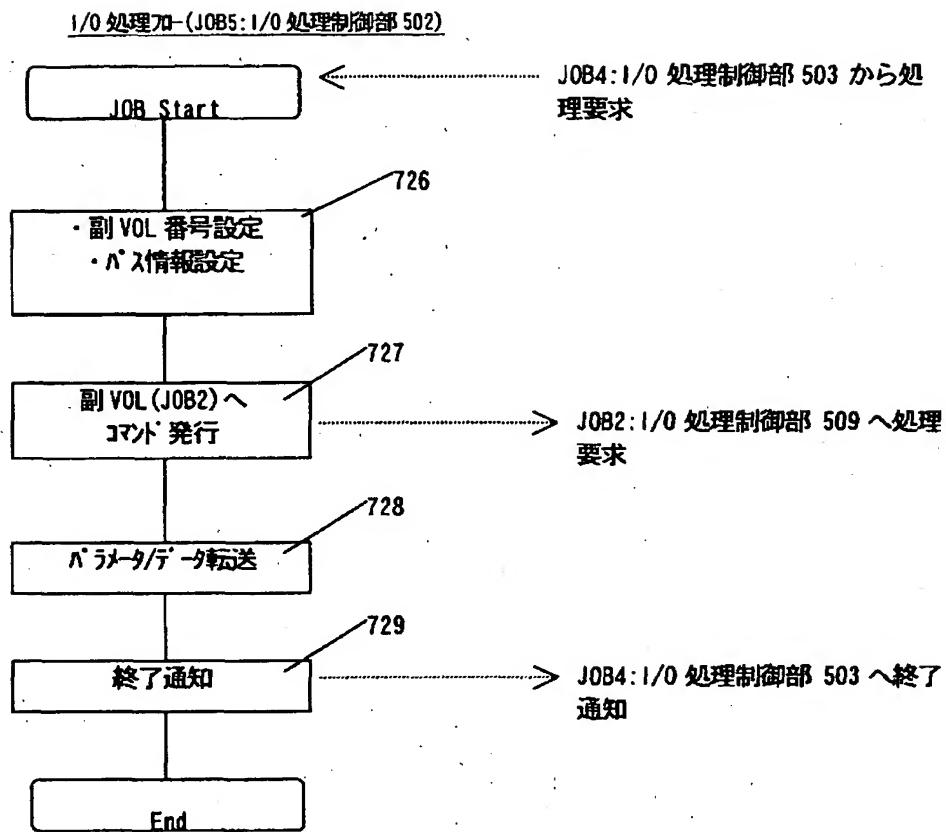
【図6】

図6

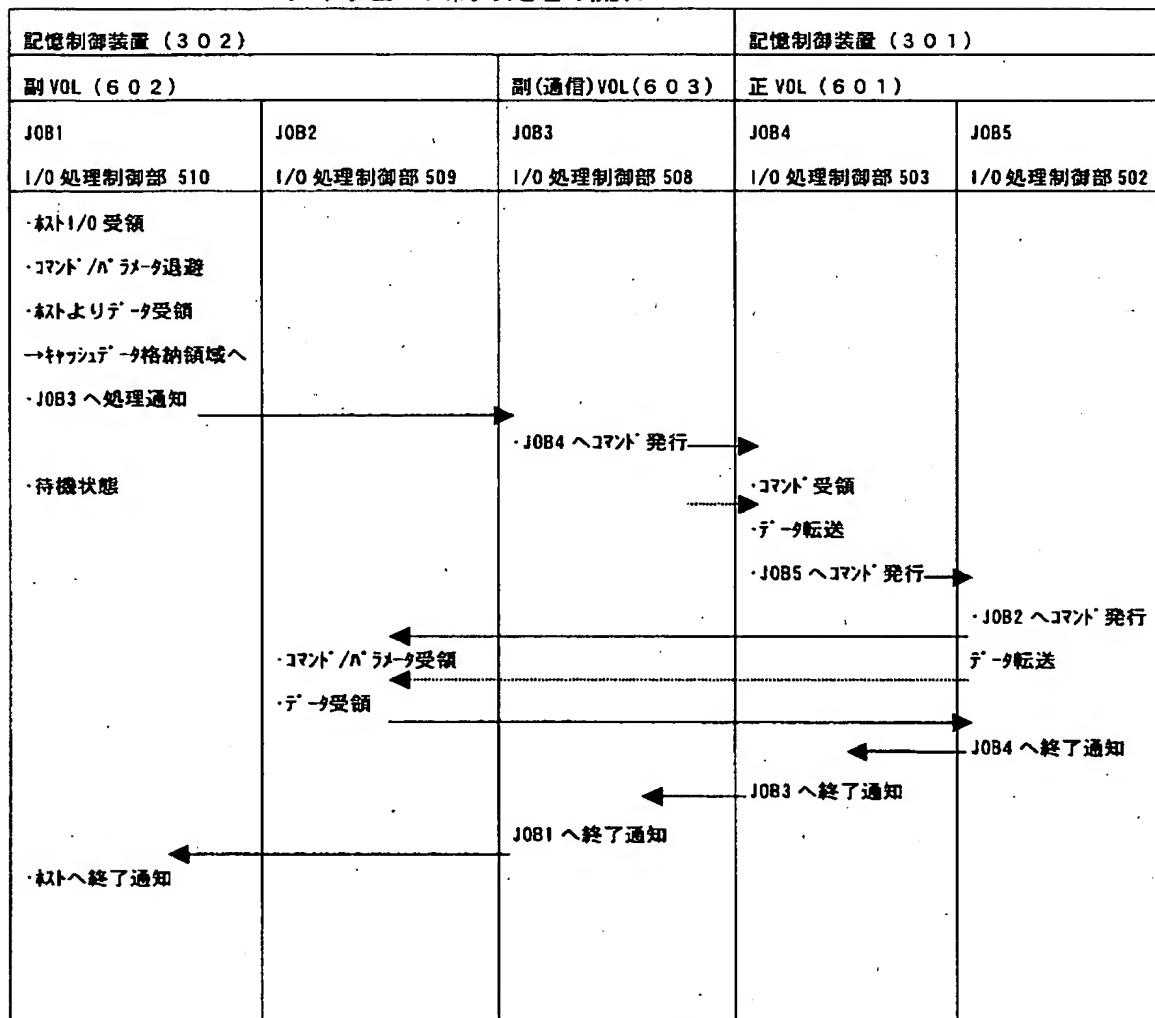


【図7】

図7

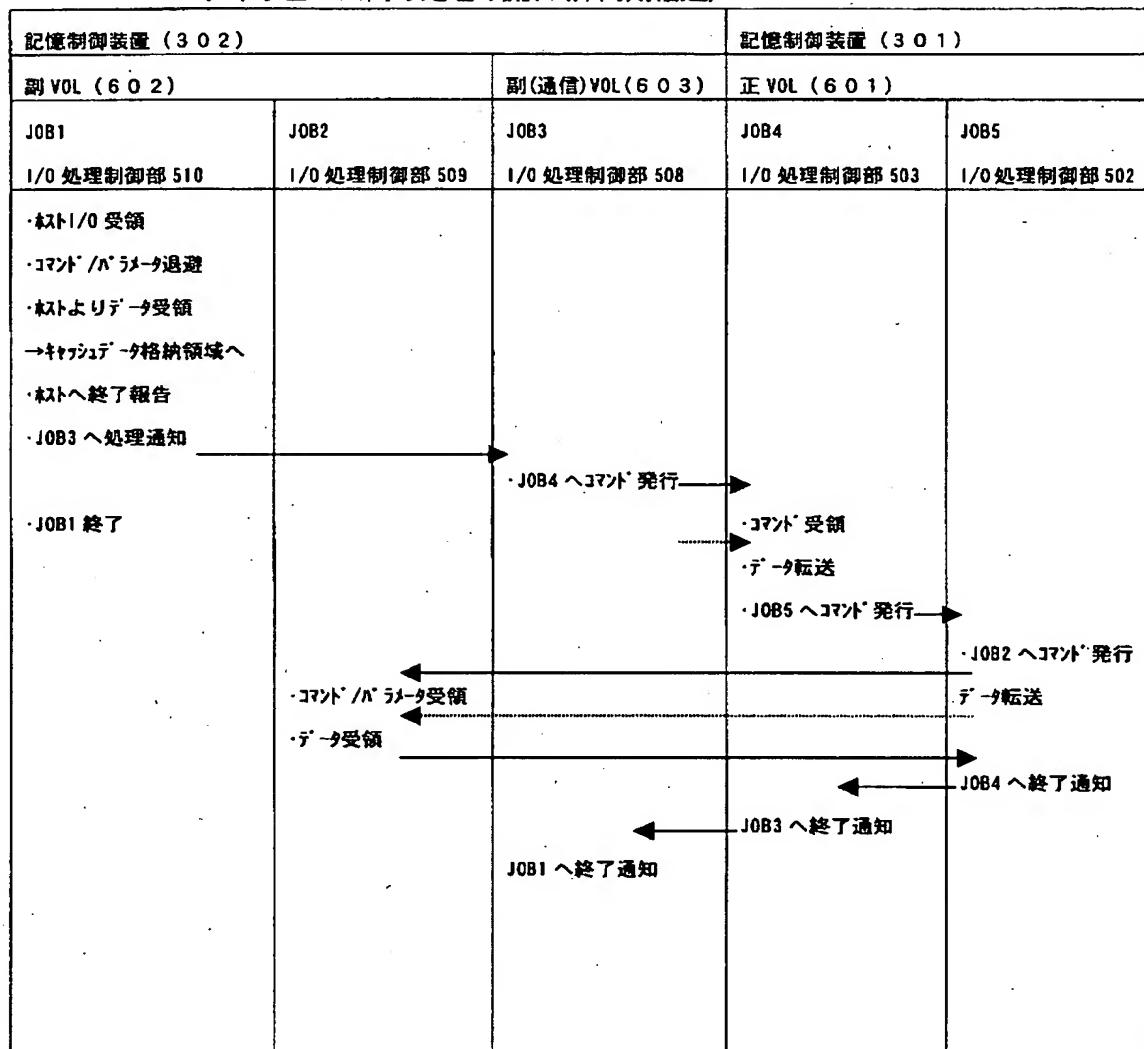


【図8】

図8
ペアボリューム間の処理の流れ

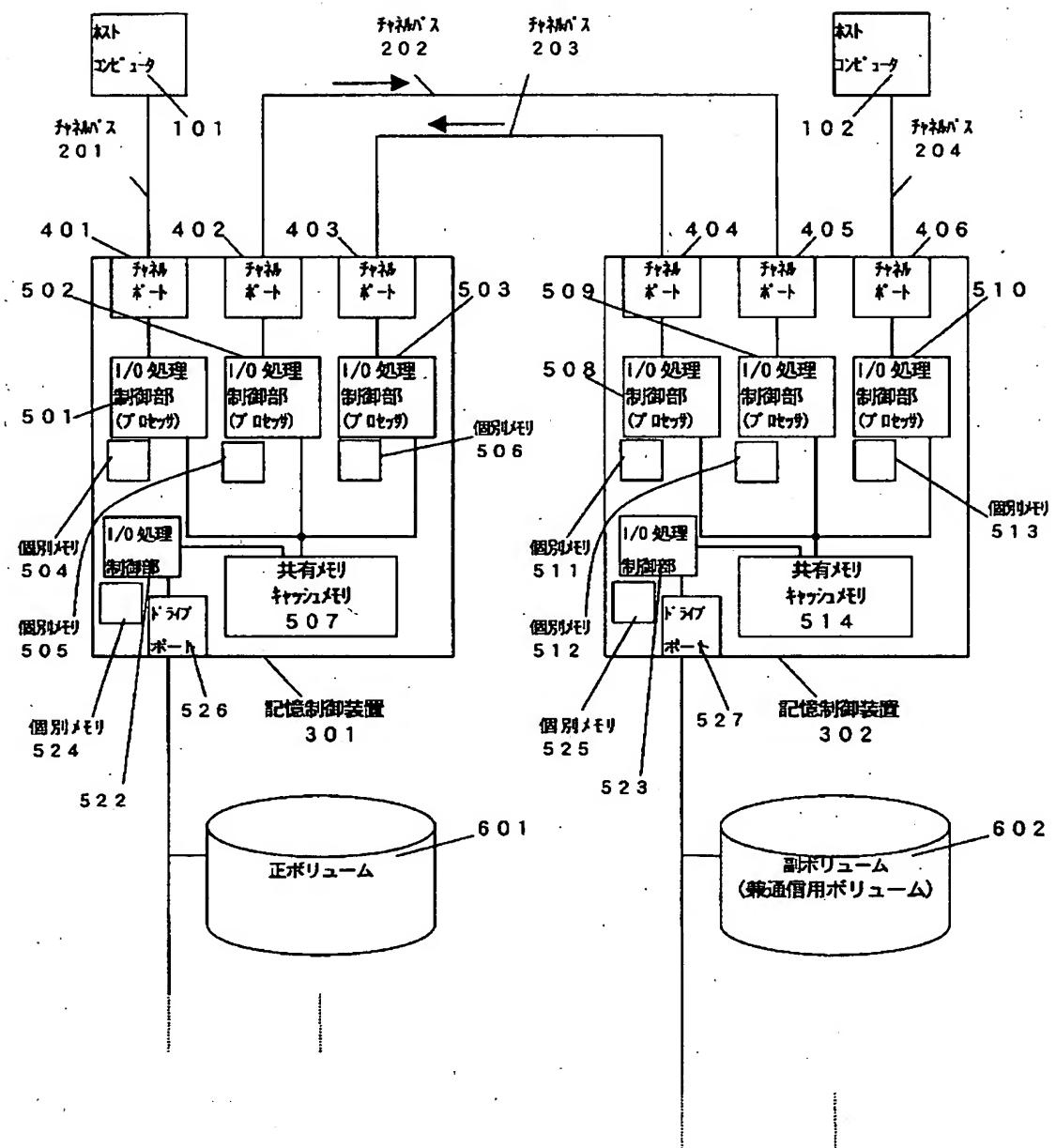
【図9】

図9
ペアボリューム間の処理の流れ(非同期転送)



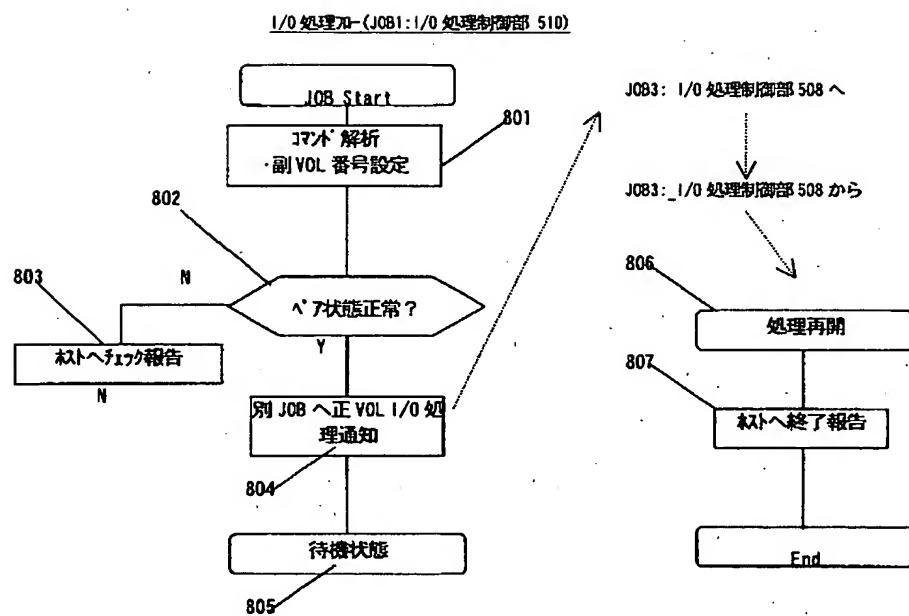
【図10】

図10



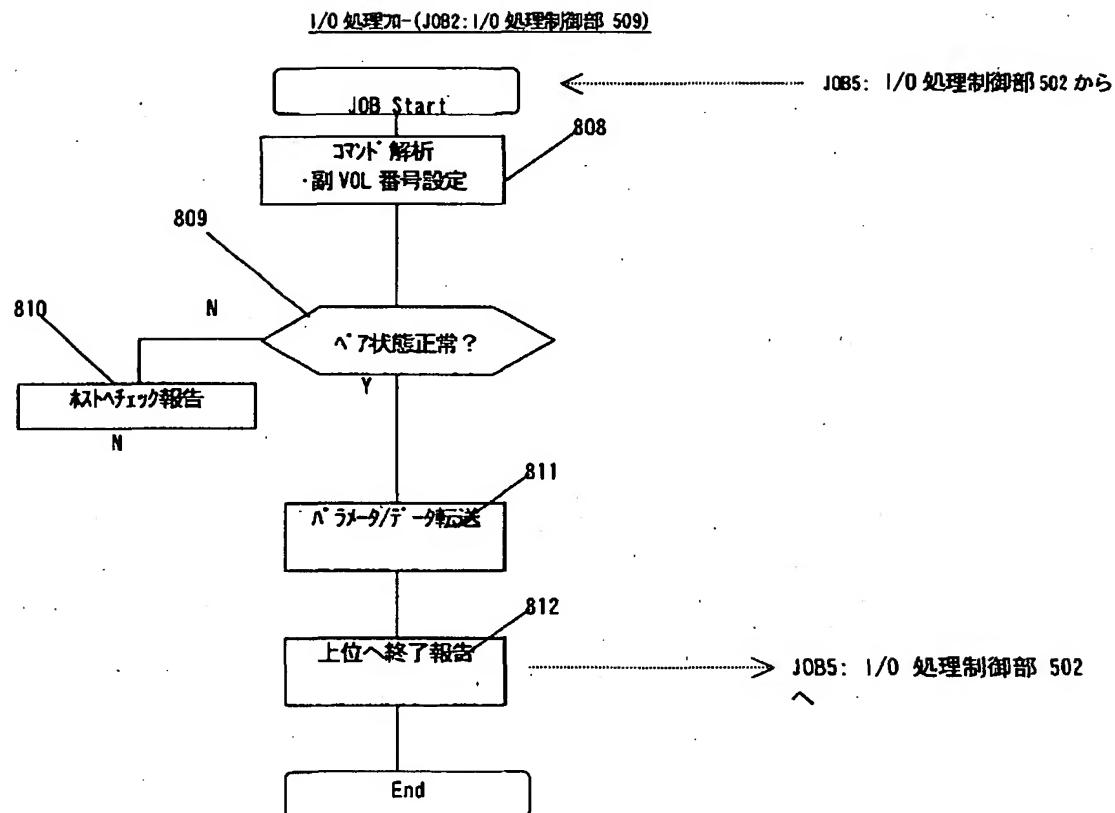
【図11】

図11



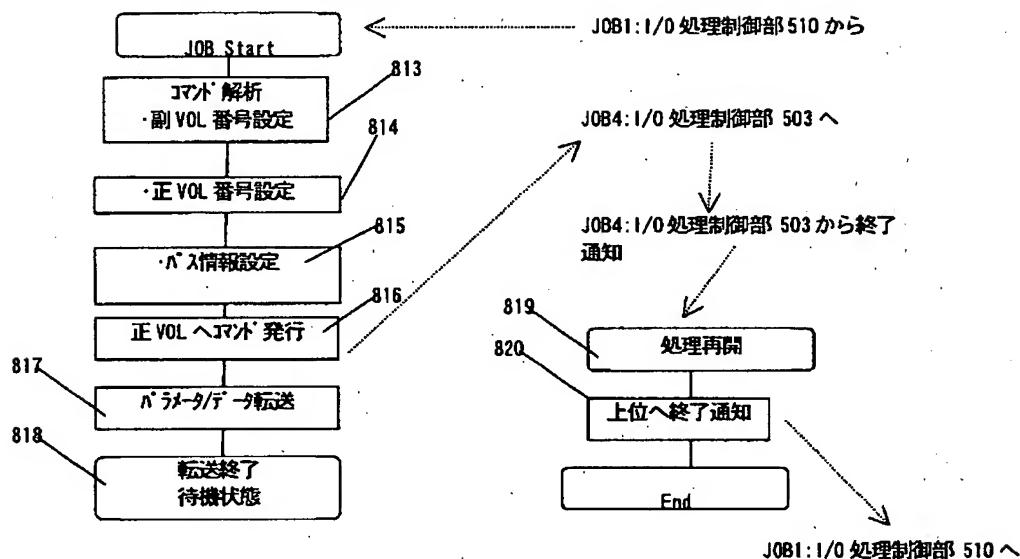
【図12】

図12



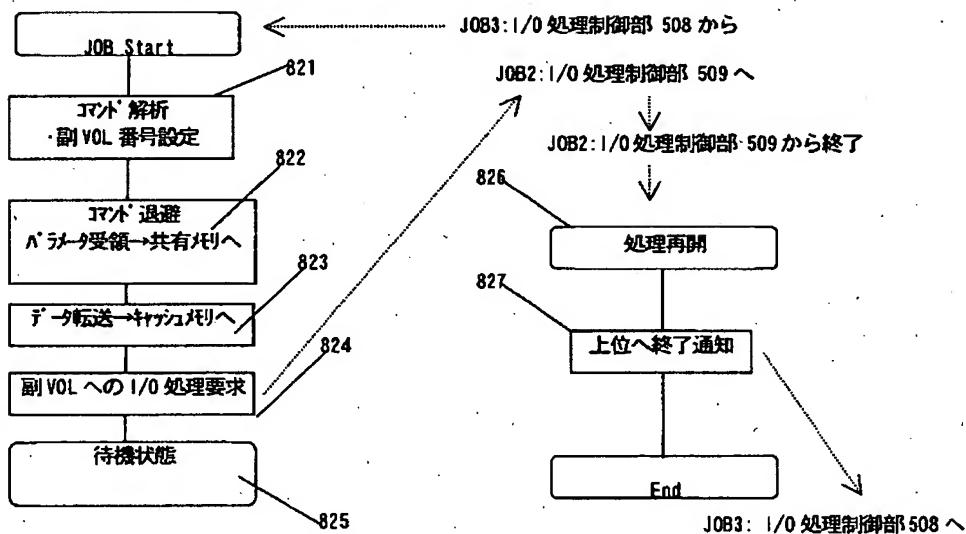
【図13】

図13

I/O処理70-(JOB3:I/O処理制御部508)

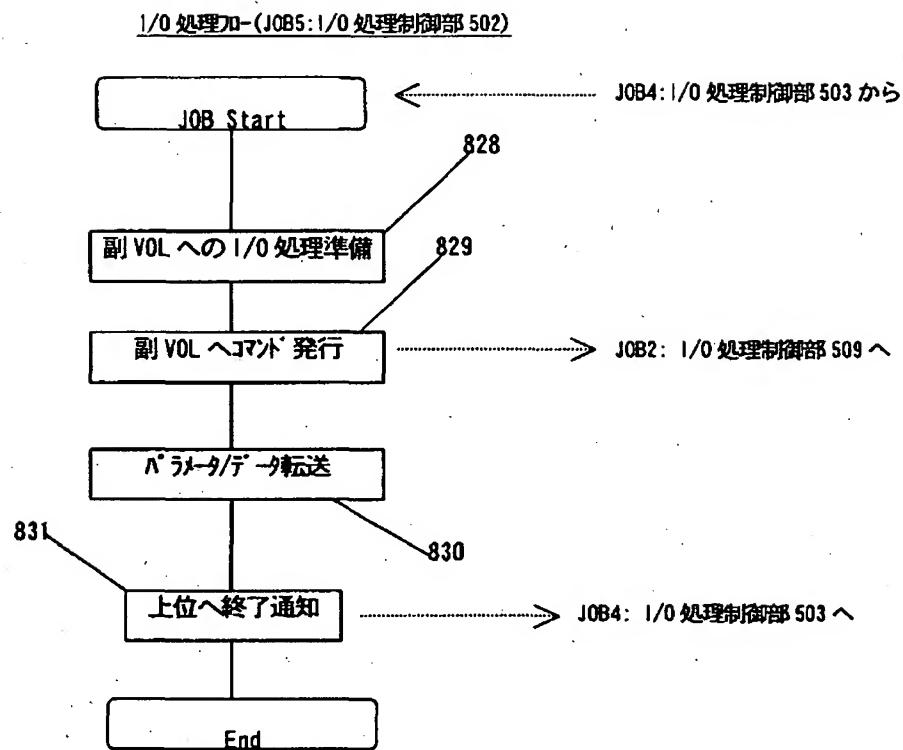
【図14】

図14

I/O処理71-(JOB4:I/O処理制御部503)

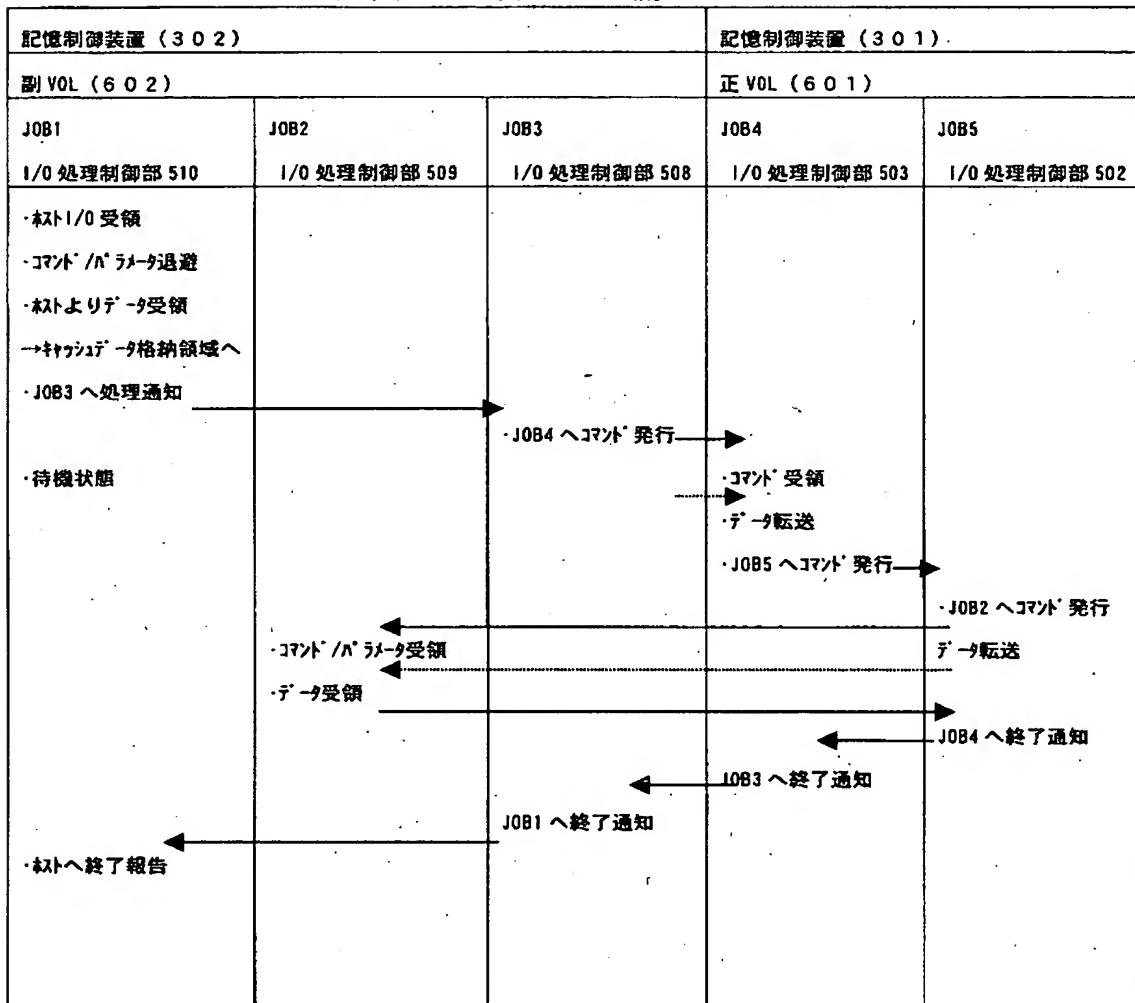
【図15】

図15



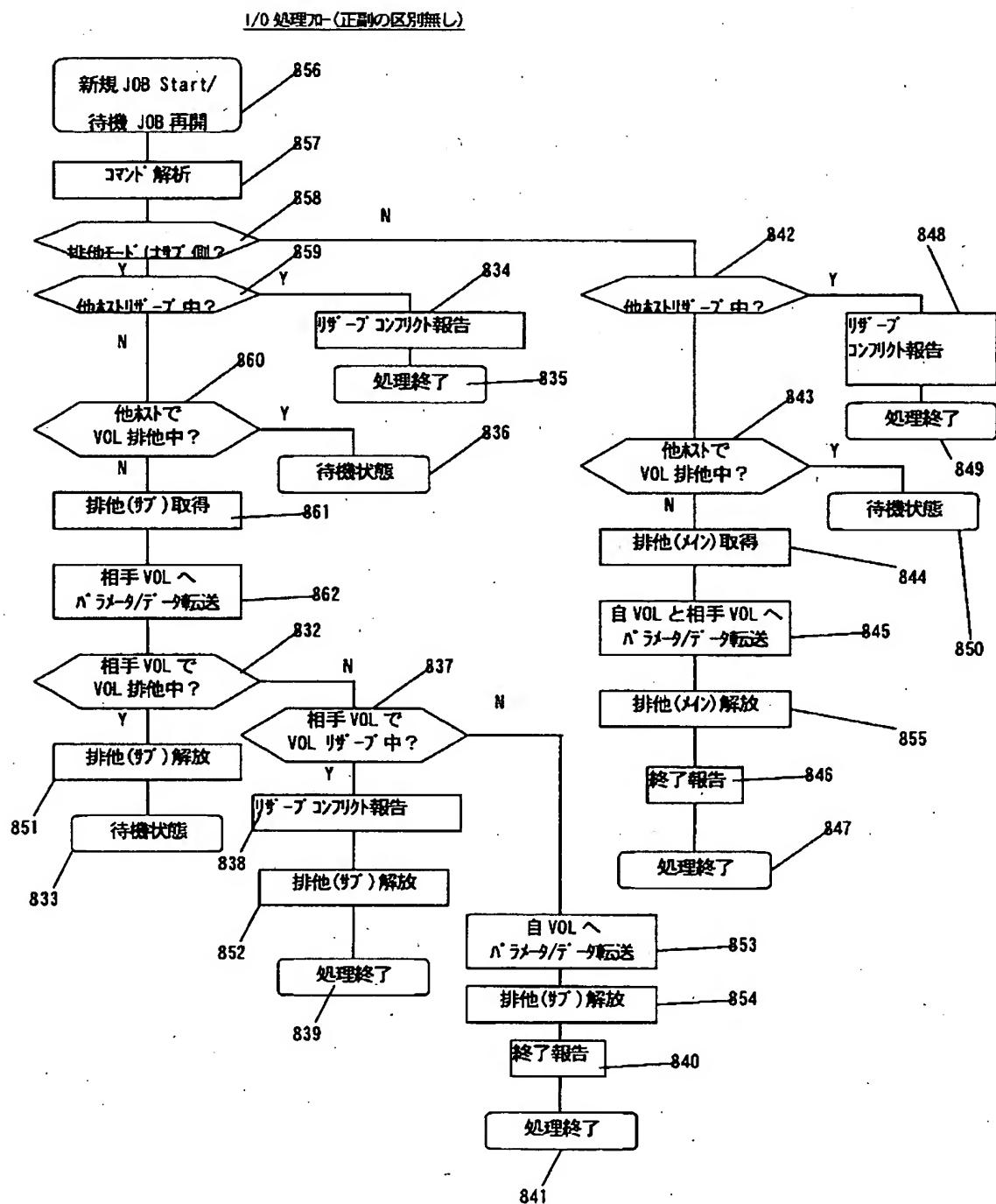
【図16】

図16 ペアボリューム間の処理の流れ



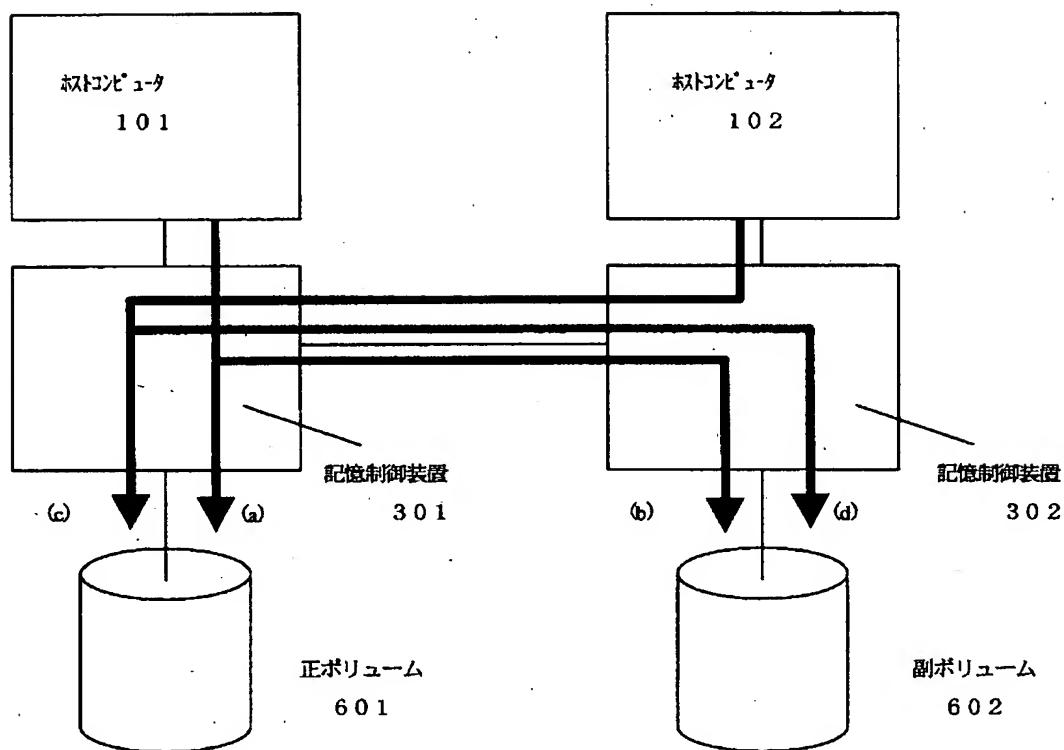
【図17】

図17



【図18】

図18



【書類名】 要約書

【要約】

【課題】 正ボリュームと副ボリュームとに格納されているデータの整合性を保持しつつ、正ホストコンピュータ及び副ホストコンピュータからの書き込み要求を受け付ける。

【解決手段】 記憶制御装置302は副ボリュームに発行されたI/O処理を、通信ボリュームを介して、正ボリュームへ転送する。正記憶制御装置301は、正ホストコンピュータから受け取った書き込みデータを正ボリューム601に書き込むと共に、副記憶制御装置302へ送る。

【選択図】 図1

認定・付加情報

特許出願の番号	特願2002-326257
受付番号	50201695286
書類名	特許願
担当官	第七担当上席 0096
作成日	平成14年11月12日

<認定情報・付加情報>

【提出日】 平成14年11月11日

次頁無

出願人履歴情報

識別番号 [000005108]

1. 変更年月日 1990年 8月31日

[変更理由] 新規登録

住 所 東京都千代田区神田駿河台4丁目6番地

氏 名 株式会社日立製作所